

国家科学技术学术著作出版基金资助出版

“十四五”时期国家重点出版物出版专项规划·重大出版工程规划项目



变革性光科学与技术丛书

Theory of Optical Waveguides

光波导理论

吴重庆 著

清华大学出版社
北京

内 容 简 介

随着光纤通信、光纤传感以及光无源器件和有源器件的飞速发展,作为它们共同的理论基础——光波导理论出现了一系列新的突破,将这些新理论、新原理加以系统化,充实到原有理论体系之中,是当前一个紧迫问题。本书改变了以往光纤理论的单模光纤、多模光纤的程式,以光波导的结构演变为主线,强调理论自身的逻辑性,系统地研究了不同结构光波导的不同概念、特性及相互联系,特别加强了光波导的模式理论、传输特性、双折射现象和模式耦合知识。全书思路清晰,知识结构完整,注重讲述不同类波导的独特分析方法,密切结合光纤通信的最新发展。

本书可供研究生及从事光纤通信、光纤传感以及波导设计的人员参考。

版权所有,侵权必究。举报电话:010-62782989, beiqinquan@tup.tsinghua.edu.cn

图书在版编目(CIP)数据

光波导理论/吴重庆著. —北京:清华大学出版社,2023.4
(变革性光科学与技术丛书)
ISBN 978-7-302-63244-3

I. ①光… II. ①吴… III. ①光波导—波导理论 IV. ①TN252

中国国家版本馆 CIP 数据核字(2023)第 057271 号

责任编辑:鲁永芳
封面设计:意匠文化·丁奔亮
责任校对:欧 洋
责任印制:宋 林

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-83470000

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:三河市铭诚印务有限公司

经 销:全国新华书店

开 本:170mm×240mm 印 张:42.25

字 数:849 千字

版 次:2023 年 6 月第 1 版

印 次:2023 年 6 月第 1 次印刷

定 价:289.00 元

产品编号:083538-01

前 言

距第2版的发行,匆匆已过十余年。在这十年中,光纤通信,光纤传感以及激光技术都在飞速发展,光波导的一些新概念、新结构、新原理、新方法纷纷涌现,如自由空间的异形波传播、四元数偏振光学、周期结构光波导(俗称光子晶体和光子晶体光纤)、保椭圆光纤、光纤应力下的双折射、电光效应、磁光效应等双折射效应、少模光纤、各向异性光波导、表面等离子波、亚波长光纤以及硅基波导,等等。在光的使用波段上,也大为拓展,如太赫兹波、紫外光波等。光波导不仅作为一种传输介质,比如光纤通信用的光纤,而且作为一种感知环境变化的传感元件,如光纤传感器,更多的是作为光信号处理器件,如光调制器等。作为它们共同的理论基础——光波导理论必须顺应这种发展潮流,将这些新理论、新原理加以系统化,充实到原有理论体系之中。同时,对于理论发展过程中的模糊概念也有必要清理和规范化。

另一方面,在吸纳这些新成果的同时,更新后的光波导理论仍然需要坚守理论自身的逻辑性,强调以光波导的结构演变为主线,坚持对于不同结构光波导理论的系统性,坚持研究突出不同概念、特性及相互联系。作为一种基础理论,光波导理论不应该被应用场合、制造工艺所左右。比如,光波导理论并没有限定使用波长,只要满足麦克斯韦方程规定的约束条件,都是正确的,所以,从理论上说,并不区分可见光波导、红外波导、紫外波导以及太赫兹波导等;再如,所谓少模光纤、亚波长光纤等,都仅仅是模式理论的一种特殊应用而已,尽管在实际的具体运用中,它们之间的区别还是很大的,然而在理论上却没有本质的不同。因此,它们都没有必要单独列出,只是在涉及这种应用状态时,增加了相关的内容。

第3版仍然特别加强了光波导的模式理论、传输特性、双折射现象和模式耦合等基本概念的讲述。与第2版相比,增加了复矢量光波、四元数偏振光学、多层平面光波导、矩形波导、周期结构圆光波导(光子晶体光纤)、渐变折射率平面光波导、波导光栅、高双折射光纤、各向异性光波导、各种感应双折射效应(弹光、电光、磁光等)、光纤的出射光场以及与自聚焦透镜的耦合等前沿问题的章节,而且从频域与时域两个方面进行阐述。

全书思路清晰,知识体系更加完备,注重讲述不同类波导的独特分析方法。鉴

于光波导理论的公式繁多,给读者阅读带来不便,本书特别加强了文字与附图的表述,使公式“活”起来。同时采用更高等的数学工具(第 12 章),使概念一目了然。希望读者能够借助这些数学工具,更深刻地理解相关的物理概念与解决问题的思路。最后,本书还密切结合光纤通信、光纤传感以及激光技术的最新发展,供研究生及从事光纤通信、光纤传感以及激光技术中波导设计的人员参考。

本书彩图请扫二维码观看。



作者

2021 年 2 月

目 录

写在前面	1
绪论	3
第 1 章 光波的一般理论	7
1.1 光场	7
1.1.1 光场的频域描述	8
1.1.2 频域麦克斯韦方程与亥姆霍兹方程	13
1.1.3 时域复振幅与复矢量的概念	20
1.1.4 时域麦克斯韦方程与亥姆霍兹方程	24
1.2 光波	28
1.2.1 光波的概念——频域描述	28
1.2.2 光波的概念——时域描述	34
1.2.3 光波纵向分量与横向分量的关系	40
1.2.4 光线的概念	44
第 1 章小结	49
第 1 章思考题	55
第 2 章 正规光波导	57
2.1 正规光波导的模式	57
2.1.1 模式的概念	58
2.1.2 模式场的纵向分量与横向分量的关系	62
2.1.3 模式的分类	64
2.1.4 正向模与反向模的关系	65
2.1.5 模式的正交性	66
2.2 传输常数的积分表达式	68

2.3 偏振问题——四元数方法	70
2.3.1 引言	70
2.3.2 偏振态描述的一般方法	72
2.3.3 偏振态描述的四元数方法	75
第2章小结	79
第2章思考题	80
第3章 均匀光波导	83
3.1 概述	83
3.2 阶跃平面光波导	86
3.2.1 模式场	87
3.2.2 特征方程	90
3.2.3 截止条件,单模传输及远离截止频率的情形	92
3.3 多层平面光波导	94
3.3.1 一般概念	95
3.3.2 转移矩阵法	96
3.3.3 多层对称平面光波导	100
3.4 矩形光波导	106
3.4.1 马卡梯里近似解法	107
3.4.2 有效折射率法	116
3.4.3 微扰法	119
3.5 圆均匀光波导	123
3.5.1 圆均匀光波导概述	123
3.5.2 线偏振模与标量法	128
3.6 二层圆均匀光波导(阶跃光纤)	131
3.6.1 二层圆均匀光波导概述	131
3.6.2 矢量法	132
3.6.3 标量法	151
3.7 多层圆均匀光波导	159
3.7.1 矢量法	160
3.7.2 标量法	162
3.7.3 三层圆光波导	165
第3章小结	168

第 3 章 思考题	169
第 4 章 正规光波导的传输特性	173
4.1 概述	173
4.2 群时延	178
4.3 脉冲展宽与色散	181
4.3.1 高斯脉冲展宽	181
4.3.2 群相移	187
4.3.3 输入非归零码的情形	188
4.3.4 更一般的脉冲展宽的概念	192
4.3.5 色散的定量描述	194
4.3.6 色散对于传输系统速率的限制	196
4.4 高阶色散与基本传输方程	199
4.5 二层圆光波导的传输特性(单模光纤的传输特性)	201
4.6 多层圆均匀光波导的传输特性	207
4.6.1 基于归一化传输常数 h 的分析方法	208
4.6.2 基于矢量法的分析方法	210
第 4 章小结	212
第 4 章思考题	212
第 5 章 非均匀光波导	215
5.1 概述	215
5.2 渐变折射率平面光波导	217
5.2.1 非均匀平面波导的射线分析法	217
5.2.2 渐变型折射率平面光波导的模式分析法	222
5.3 非均匀平面光波导的近似解法	226
5.3.1 WKB 近似法	227
5.3.2 变分法	231
5.3.3 多层分割法	238
5.4 平方律圆光波导	239
5.4.1 射线法	240
5.4.2 非均匀圆光波导的模式解法	247

5.4.3	直角坐标系法	249
5.4.4	柱坐标系法	251
5.5	非均匀圆波导的近似解法	256
5.5.1	高斯近似法	256
5.5.2	级数解法	261
5.5.3	伽辽金方法(变分法)	266
5.5.4	少模与多模光纤的模式问题	270
第5章小结		277
第5章思考题		278
第6章 非正规光波导		280
6.1	概述	280
6.2	正规光波导的辐射模与空间过渡态	284
6.3	纵向模耦合方程	286
6.3.1	弱导缓变光波导的模式耦合	286
6.3.2	矢量模耦合方程	289
6.4	光波导光栅	293
6.4.1	均匀光栅的一般原理	294
6.4.2	光纤光栅	299
6.5	突变光波导	303
6.5.1	光纤对接	304
6.5.2	光纤的出射光场	308
6.5.3	光纤与自聚焦透镜的耦合	314
第6章小结		318
第6章思考题		320
第7章 非圆光波导		322
7.1	非圆光波导与双折射	322
7.1.1	非圆光波导的概念	322
7.1.2	双折射现象	324
7.2	非圆均匀光波导	326
7.2.1	一般解法(微扰法)	326

7.2.2	二层非圆光波导	329
7.2.3	椭圆芯阶跃光纤	330
7.2.4	非圆非均匀光波导	333
7.3	偏振态演化的概念	338
7.3.1	概述	338
7.3.2	无损光波导的传输矩阵	340
7.3.3	缪勒矩阵	342
7.3.4	偏振主轴	347
7.3.5	双折射分析的四元数方法	349
7.4	任意坐标系下非圆正规光波导的一般解	356
7.5	非圆非正规光波导中偏振态的演化	364
7.5.1	偏振模耦合	364
7.5.2	非圆非正规光波导的偏振特性	370
7.6	偏振模色散	378
7.6.1	偏振模色散的概念	378
7.6.2	偏振模色散对光信号脉冲波形的影响	388
7.7	高双折射光纤(保偏光纤)	393
7.7.1	消光比与串音	394
7.7.2	保偏光纤稳定输出偏振态的原理	402
7.7.3	线偏振态保持光纤和椭圆偏振态保持光纤	405
第7章小结		410
第7章思考题		412
第8章	微结构圆光波导	416
8.1	微结构圆光波导的概念	416
8.2	内全反射型微结构圆光波导	419
8.2.1	模式场和它满足的方程	420
8.2.2	圆对称化	423
8.2.3	其他方法简介	426
8.3	带隙光纤	428
8.3.1	概述	428
8.3.2	空间周期性函数的频域展开	430
8.3.3	本征值方程	436

8.3.4 二值二维空间周期函数的频谱	442
第 8 章小结	448
第 8 章思考题	449
第 9 章 各向异性光波导	451
9.1 各向异性光波导的概念	451
9.1.1 材料的各向异性	451
9.1.2 光在均匀各向异性材料中的传播	455
9.1.3 光从各向同性介质入射到各向异性介质时界面上的 双折射	464
9.2 应力双折射	469
9.2.1 本征双折射与感应双折射的概念	469
9.2.2 弹光效应与应力双折射	470
9.2.3 应力型保偏光纤外应力作用下的偏振性能	480
9.2.4 利用四元数方法分析多重双折射问题	491
9.3 电光效应双折射	498
9.3.1 介质的电光效应	498
9.3.2 各向同性介质的电光效应	508
9.3.3 各向异性介质的电光效应	516
9.4 旋光性和磁光效应	522
9.4.1 旋光性	522
9.4.2 磁光效应	527
9.4.3 隔离器与环行器	531
第 9 章小结	532
第 9 章思考题	534
第 10 章 光波导间的横向耦合	536
10.1 正规光波导的横向模式耦合	537
10.1.1 横向模耦合方程	537
10.1.2 耦合波的特性	542
10.1.3 用极化电流概念推导模耦合系数	544
10.1.4 两正圆光纤的横向耦合	546
10.2 缓变非正规型横向耦合	550
10.3 光纤耦合器	554

10.3.1	2×2 光纤耦合器	555
10.3.2	3×3 光纤耦合器	559
10.4	光纤干涉仪	567
10.4.1	光纤马赫-曾德尔干涉仪	567
10.4.2	光纤迈克耳孙干涉仪和斐索干涉仪	570
10.4.3	无源光纤环	577
10.4.4	太赫兹光学非对称解复用器	588
10.4.5	附录:反演矩阵与法拉第旋镜	591
10.5	基于 3×3 耦合器的干涉仪	593
10.5.1	光纤环形谐振腔——频域	593
10.5.2	3×3 耦合器干涉仪的时域应用	598
	第 10 章小结	603
	第 10 章思考题	605
第 11 章	光波导的损耗与增益	606
11.1	光波导损耗的一般概念	607
11.1.1	无限大均匀介质的损耗	607
11.1.2	光波导的波导损耗	613
11.2	光纤的传输损耗	618
11.2.1	石英光纤的材料损耗	619
11.2.2	光纤的波导损耗	622
11.3	光纤损耗的测量与光时域反射技术	627
	第 11 章小结	632
	第 11 章思考题	634
第 12 章	相关数学公式	636
12.1	复矢量	636
12.2	四元数代数	645
第 13 章	总结与汇总表	651
13.1	总结	651
13.2	各类光波导汇总表	656
	参考文献	661

清华大学出版社

写在前面

一提到理论,首先想到的便是一堆眼花缭乱的公式;理论书籍给人的印象,就是公式推过来推过去。老师上课也是推一黑板的公式,不厌其烦。

然而,当今计算机模拟仿真与计算技术已经相当发达,各种光学问题都可以用计算机求解,设计及模拟实验,公式推导与演算也完全可以交由计算机解决,如果能省去公式推演的理论课,岂不美哉?那我们要问,在计算机高度发达的环境中,以公式推演为主的理论还有什么用?换言之,我们为什么还要学习像“光波导理论”这样的理论课?

以公式和数学符号为特征的理论课,大概有如下三个作用:

(1) 用于计算。就是在做一个实验之前,先计算一下,预测一下结果。或者在实验之后,看看是否与理论计算相一致,用来检验实验的正确性。这个作用已经被计算机所代替。

(2) 用来精确而全面地描述现象,将现象抽象为物理量并最终用一个数学符号描述。比如电现象,也就是电荷相互作用的现象,首先将这个现象解释为电场,并由此提炼出描述电场的量是电场强度,进一步再把电场强度归纳为一个矢量,最后这个矢量用 E 表示。长期以来,由于应试教育,只重视解题而把概念的提炼过程忽视了,造成对数学符号的反感。数学符号具有文字描述所不具有的精确性和全面性。当我们看到符号 E 时,我们立即会反映出它是一个具有大小和方向的量,这就是它的精确性。同时,如果 $E=0$,说明是一个零矢量,它的大小为零而方向为任意方向。这就是它的全面性。

(3) 对物理量的运算,可以产生新的概念。以力矩 $M=r \times F$ 为例,该等式告诉我们,力矩可以通过测量力 F 和力的作用点到原点的距离矢量 r 经过叉乘运算得到。或者说,对这两个量(力和距离)进行运算,产生了第三个量(力矩)。这种利

用数学运算产生新概念的方法,在理论课中尤为重要。比如,光波导中的模式,其实就是描述光波导传输特性的微分方程的一个个特解。到目前为止,人们相信,如果概念所定义的物理量正确反映了它的特征,而且推导过程是正确的,那么它的结论就必然是正确的。因此,理论可以用来精确地预言。在这个过程中,正确的理论推导是关键,甚至一点点小的疏忽都将导致错误的结论。

因此,无论计算机技术多么发达,甚至包括人工智能,仅仅可以取代理论分析的第一个作用,也就是代替仿真和计算。而第二、第三个功能是取代不了的。至少目前还是这样。数值计算只能是有限的特例,也可能漏掉一些特殊现象。为了获得新概念、预示新现象,找出不同概念的联系,区分同种效应不同现象之间的分界,必要的理论课学习仍然是不可或缺的。

正因为理论课的作用已经从计算转移到概念的全面描述和物理量关联的描述,因此特别需要注意:

(1) 尽量使用先进的数学工具描述物理概念。麦克斯韦方程组是一个典型的例子。麦克斯韦方程组使用矢量,矢量微分算符(∇)以及矢量的叉乘、点乘等运算。它不仅将大量的分量运算描述的电磁场关系(十多个公式)大大简化,而且概念非常清晰。所以使用先进的数学工具容易导致理论上的成功。

(2) 看到一个数学符号,立刻应该联想到它的物理意义。例如,看到一个矢量符号 $\mathbf{E}$,立刻就想到它的大小和方向,同时也会想到它具有 3 个分量,而且它也是时间与空间的函数,因此,它相对于时间的变化率和对空间的变化率是互相关联的,等等。再加看到一个复矢量符号 $\mathbf{E} = a_x \hat{x} + a_y \hat{y}$ (a_x, a_y 均为复数),我们应该立刻联想到它是一个偏振态。如此等等。

(3) 经常将用符号和公式描述的理论,转换成它所描述的物理现象或者物理规律。也就是“用语言把公式说出来”。比如,看到公式 $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$,我们应该立刻能用语言说出来,即电场强度随空间变化率的旋度分量正比于磁场强度随时间的变化率,二者方向相反。或者说,磁场的时变化引起电场的空间变化。

(4) 最后一点,要注意那些公式的用途,在实际技术应用中有什么指导意义。如果在学习中的关注点从数值计算、解题等转移到以上 4 点,就不会再觉得理论枯燥,也就有心思坚持读下去了。希望读者牢记这一点,本书作者也力争做到这一点。

另外,本书为了保持其严谨性,有大量的理论证明。如果读者不怀疑理论的正确性,这部分内容可以忽略不看,只注重结论的应用就可以了。

绪 论

光纤通信、光纤传感以及激光技术的发展,推动人类社会向信息社会的变革。1970 年第一根低损耗光导纤维的出现,翻开了人类通向信息社会新的一页。研究光如何在光纤和各种光波导中传输的理论,即光波导理论,则是光在不同波导结构的介质中传输的基本理论,它指导着光纤技术与激光技术的前进。

虽然在古代人们对光的传播就有了一定的认识,但深入研究光在光纤和各种光波导中的传播理论,却是近五十年的事。光波导理论源于微波波导理论。20 世纪 50 年代后期,电子学的发展,使人类对电磁波的利用,推进到了微波波段,全世界都在致力于微波波导的研究。然而历史却没有遵从“循序渐进”的原则,当微波技术远没有获得所预料的广泛的大规模应用的时候,光波的时代就来到了。

从微波到光,是人类对电磁波利用的必然趋势。所以光波导理论和微波波导理论有密切的联系,然而二者却有截然不同的特点。光波导不只是介质微波波导尺寸的缩小,光在光纤中的损耗机理,光波导的弱导性及其他传输特性都与微波波导不相同,所以光波导理论是一门独立的理论。

人类对光本质的认识经历了曲折的历程。如今,因光在传输过程中表现出波动性而将光看作一种电磁波已成为人们的共识。虽然在某些场合,利用几何光学或光线理论可以得到直观的感性认识,但却是肤浅的;而光的量子性在光传输过程中表现甚微,常被忽略。本书正是将光看作一种电磁波来研究光波导的。

光波既然是电磁波,那么它首先是一种电磁振荡。电磁振荡包括电场与磁场两方面的振荡,因此,光场按工程惯例以电场强度 $\boldsymbol{E}$ 和磁场强度 $\boldsymbol{H}$ 来表征,可写成

$$\begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix} = \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix} (x, y, z, t)$$

光场既是位置 $\boldsymbol{r}(x, y, z)$ 的函数,又是时间 t 的函数。

一个单一频率的简谐电磁振荡,通常可表示为

$$\begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} (x, y, z) e^{-i\omega_0 t}$$

式中, $\mathbf{E}$ 和 $\mathbf{H}$ 是复矢量,包括方向、幅度和相位关系。一个理想的相干单色光,也可以写成上述形式,即光场中每点的场分量均有固定的方向,稳定不变的频率和稳定的相位关系。但它的频率极高,达到 10^{14} Hz 量级。即使延伸到太赫兹光,它的振荡频率也在 10^{13} Hz 以上。然而到目前为止,我们仍然没有可以直接观测到光场振荡的技术,也就是说光是一种电磁波的基本假说仍然停留在间接证明的阶段,还没有直接观察到电磁振荡。实际的光场是由许多这样理想的相干光场叠加而成,包括连续的或离散的叠加。它们不仅可能频率不同,而且亦无稳定的相位关系,方向也不固定。所谓自然光就是一种这样的光。

以上是对光的一个简单而基本的认识。下面再分析光波导。

光波导是一种具有明确界面和确定折射率分布的传输介质,它对光的传播具有确定的限定条件。大部分光波导都是人造的,虽然大气、海洋也是一种传输介质,但人们并不称它们为光波导,因为它们的界面不明确。具有纵向与横向的取向区分是光波导最重要的限定条件,并以此作为一个波导的特征。纵向往往定义为传输方向,横向定义为垂直于纵向的横截面。然而,这种取向区分完全是人为的,具有任意性,我们以光在无限大平面的界面的三种媒质中的传播为例说明这个问题。

如果以图 0-1(a)中 x 方向为纵向,我们看到光是从介质 n_1 经过介质 n_2 传播的。在界面上服从折射定律和反射定律。但如果我们以图 0-1(b)的 z 方向为纵向,则它是一个三层平面光波导。若 $n_2 > (n_1, n_3)$, 光可束缚在 n_2 介质内传播,它在 x 方向上呈现一个稳定的场分布(模式)。显然,这种场分布和前面的折射定律和反射定律所得的结果是一致的。由此可见,对纵向的不同定义,决定了对光波导特性的不同描述。不过,既然已经把它看作一种波导而不看作光场(电磁场)存在的一般介质空间,总是要规定纵向的。

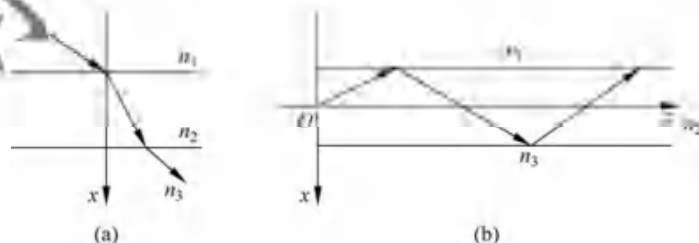
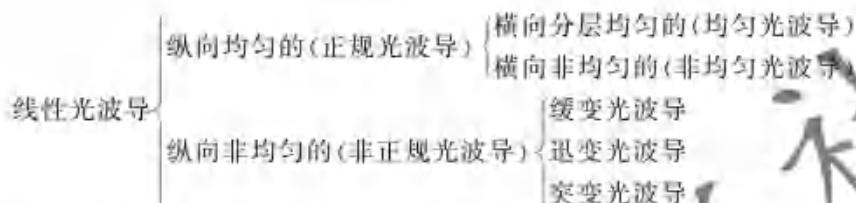


图 0-1 光波导取向区分

光导纤维是一种最重要、最常见的圆光波导。通常,光导纤维的纵向长度比横向尺度大得多,材料折射率的分布沿纵向基本是均匀的,或在局部可看成均匀的。光波导是比光导纤维更为广阔的概念。

根据光波导折射率空间分布的均匀性,我们进行如下的分类:



这种分类完全是从便于理论分析的角度出发的,根据需要,实际的光波导或者光纤可看作其中的某一类。比如当只考虑某根光纤的基本模式时,可把它看作纵向均匀的正规波导。而当考虑这些模式耦合的时候,可以把它看作纵向非均匀的非正规波导。

进一步,还可以对上述分类加以细化。从光波导的制造技术来看,基本上可以分为两类,一类是基于拉制技术的圆光波导——光纤;另一类是基于刻蚀技术的光波导,有些文献中的光波导就特指刻蚀的光波导。刻蚀的光波导具有平面的界面,所以又称为平面光波导、矩形波导、脊波导等。为了使光波导理论不受制造技术发展水平的限制,在本书的同一章中,这两类波导都会研究,请读者自己加以区分。当然,在光纤上也可以刻蚀,那就另当别论了。

影响光波导传输特性的,主要是折射率沿着波导的空间分布。大量光波导的折射率分布是线性和时不变的,而且是无源的,即

$$n = n(x, y, z)$$

它与时间 t 、光场 $\begin{pmatrix} E \\ H \end{pmatrix}$ 均无关。在本书中,如果不加特殊说明,则默认服从三个假设:①线性的;②时不变的;③无源的。

需要说明的是,近年来各向异性材料被广泛应用于制作光波导,而各向异性介质的折射率是用折射率张量描述的。折射率张量来自介电张量,它是一个 3×3 的矩阵,可表示为

$$\epsilon = \epsilon_0 \begin{bmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz} \\ \epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz} \end{bmatrix}$$

一提起张量,对于没有学过张量的读者,就会觉得很抽象、很晕。其实,我们不必过多地深究它,只要把它看作一个矩阵就可以了。它导致的后果是,电位移矢量 $\mathbf{D}$ 与电场强度 $\mathbf{E}$ 不再同方向。各向异性材料仍然属于线性介质,也可以看作时不

变的。

有三种情形突破上述限制：①非线性；②时变性；③有源。非线性是指光波导的性质与输入光的光强大小有关的现象。在强激光注入或者当一根光纤中注入很多波长的光信号时，这种现象会显现出来。注意，并不是说在弱光时就没有非线性，而是说它表现不明显，所谓非线性阈值是一个不准确的概念。时变性常常在光纤作为传感元件时表现出来，比如光纤水听器就是利用光纤的折射率在声场作用下折射率改变的原理制成的；再如对于铌酸锂晶体在电压作用下折射率会发生变化，这种变化常常被称为“光折变”，即光的折射率变化的意思，然而这个术语与中文的构词规律不协调，容易误解为光被折弯而变化，很别扭。不如称为时变性更为确切。有源光波导是光有源器件的基础，比如激光器、探测器以及光放大器。有源光波导涉及光子与介质粒子的相互作用。在分析有源光波导的半经典理论中，光仍然作为一种电磁波，因此本书的理论也部分适用于有源光波导。

以上是对光波导的一个总体的概括。

光波导的传输特性是指不同类型光波导用于信息或者能量传递时的最基本的特性，它包括光本身在光波导中的传播特性和载有信息的光信号的传输特性两方面。光本身在光波导中的传播特性包括：①光场的分布形式；②传播常数或相移常数；③偏振特性；④模式耦合特性，等等。载有信息的光信号的传输特性包括：①群时延和群速度；②色散特性、高斯脉冲展宽以及群相延；③偏振模色散，等等。针对不同类别的光波导，分别研究它们的这些特性，就是贯穿于本书的主线。

以上概括地说明了光波导理论就是以光的电磁波理论为基础，研究光在各类不同介质和不同结构的波导中的传输特性的理论。

第 1 章

光波的一般理论

人一睁开眼,便接触到光彩的世界。然而人类对于光本质的认识,却经历了漫长而曲折的过程。目前,光是一种电磁波,已经成为人类的共识。但在物理光学中所学的关于光波的知识较为初步,需要进一步深入。物理光学并不强调光场与光波的差别,而且仅仅学习了平面波的相关知识。本章首先对光场的概念进一步深化,导出了光场的频域描述、光场的频域麦克斯韦方程和亥姆霍兹方程;然后建立了时域光场的复振幅与复矢量光场的概念,导出了相关的麦克斯韦方程。

就光波导的理论分析方法而言,总体上可分为频域法和时域法两种。我们首先介绍基于将光信号看成一系列频率分量叠加的频域法,然后介绍时域法。本章的概念比较抽象,公式也比较抽象,但是它们是整个光波导理论的基础,还要请读者耐心读完。

1.1 光场

光场就是一种以光频振动的电磁场,即频率范围为 $10^{13} \sim 10^{17}$ Hz 频段的电磁场,如图 1.1.0-1 所示。目前,人们的视野已经把光波的范围拓展到从太赫兹波段 (10^{13} Hz) 到紫外光的很宽泛的波段,这样,频率范围就拓展为 $10^{12} \sim 10^{16}$ Hz。尽管光频电磁场具有明显的波动性,但是,弥漫于空间的光可能来自四面八方,有相干光、散射光以及各种自发辐射和受激辐射光,它们或者相干叠加(按照振幅相加),或者非相干叠加(按照功率叠加),因此,我们不能笼统地将弥漫于空间的光场都说成是光波。尤其是在等离子体表面激发出的近场,它不具有波动性,尽管它仅仅分布于激发源周边很小的范围内,但毕竟是一种场而不是波。因此,光场与光波不是一个概念,它们之间存在细微的差别,光场考虑的是光在空间一点的电磁振动,而

光波考虑的是空间不同点电磁振动的关联性。

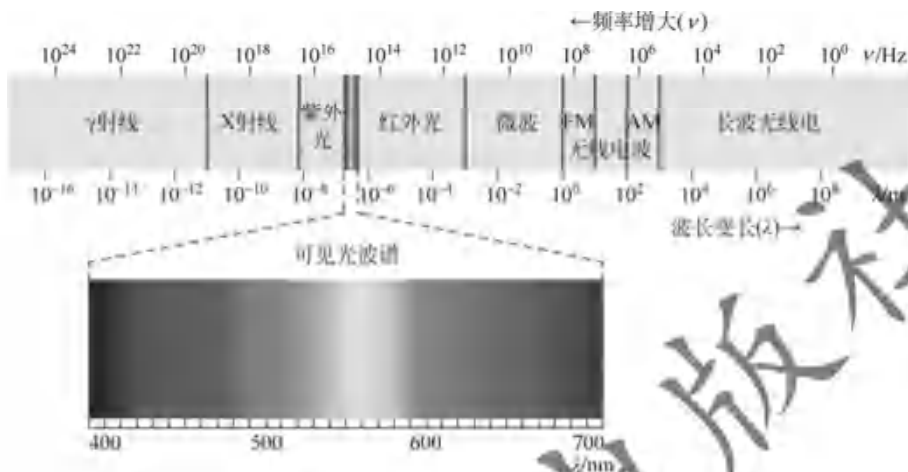


图 1.1.0-1 电磁波谱
(请扫二维码看彩图)

1.1.1 光场的频域描述

1864 年,麦克斯韦(J. C. Maxwell)回顾和总结了前人关于电磁现象的实验研究成果,提出了一套完整的宏观电磁场方程,预言了电磁波的存在并提出“光就是电磁波”的重要论断,开创了光的经典电磁理论的新纪元。迄今为止,在光通信、光集成(集成光学)、光信息处理以至于整个光学领域,有关光的问题,仍然以麦克斯韦方程作为理论基础,它推动着光波技术的发展。

描述在真空或介质中某一时刻、某一位置上电磁场的基本物理量是电场强度 $\mathbf{E}$ 和磁场强度 $\mathbf{H}$,可写成 $\mathbf{E}(x, y, z, t)$ 和 $\mathbf{H}(x, y, z, t)$,或者写成 $\mathbf{E}(\mathbf{r}, t)$ 和 $\mathbf{H}(\mathbf{r}, t)$ 。注意,此时 $\mathbf{E}$ 和 $\mathbf{H}$ 的大小都是实数。

作为以光频振动的电磁场,它们共同的基本特征是含有频率为 ω 的高频项。本节首先考虑以单一频率振动的光场,并按照直角坐标系分解为 3 个分量之和,用 $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$ 表示空间位置,其中 $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ 为坐标系 3 个方向的单位矢量,也就是

$$\begin{cases} \mathbf{E}(\mathbf{r}, t) = E_x(\mathbf{r})\cos[\omega t + \varphi_{E_x}(\mathbf{r})]\hat{\mathbf{x}} + E_y(\mathbf{r})\cos[\omega t + \varphi_{E_y}(\mathbf{r})]\hat{\mathbf{y}} + \\ \quad E_z(\mathbf{r})\cos[\omega t + \varphi_{E_z}(\mathbf{r})]\hat{\mathbf{z}} \\ \mathbf{H}(\mathbf{r}, t) = H_x(\mathbf{r})\cos[\omega t + \varphi_{H_x}(\mathbf{r})]\hat{\mathbf{x}} + H_y(\mathbf{r})\cos[\omega t + \varphi_{H_y}(\mathbf{r})]\hat{\mathbf{y}} + \\ \quad H_z(\mathbf{r})\cos[\omega t + \varphi_{H_z}(\mathbf{r})]\hat{\mathbf{z}} \end{cases} \quad (1.1.1-1)$$

对于初次学习光波导理论的人来说,上来就是这么长的公式肯定不习惯。其实,这个公式表示的只不过是一个矢量可以分解为三个分量之和;而每个分量都以同一个频率振动,每个分量的大小与相位不一定相同。也可以将上面的大长式子按照各个分量($i=x, y, z$)写成

$$\begin{cases} E_i(\mathbf{r}, t) = E_i(\mathbf{r}) \cos[\omega t + \varphi_{Ei}(\mathbf{r})] \\ H_i(\mathbf{r}, t) = H_i(\mathbf{r}) \cos[\omega t + \varphi_{Hi}(\mathbf{r})] \end{cases} \quad (1.1.1-2)$$

在这个表达式中除频率 ω 作为一个基本量不变,其余部分如它的幅度 $E_i(\mathbf{r})$ 和 $H_i(\mathbf{r})$,以及相位 $\varphi_{Ei}(\mathbf{r})$ 和 $\varphi_{Hi}(\mathbf{r})$ 都是频率 ω 的函数。在式(1.1.1-2)两个表达式中,既含有频率 ω ,也含有时间 t ,因此不能说它是时域的表达式还是频域的表达式。

需要说明的是, $\varphi_{Ei}(\mathbf{r}, t)$ 和 $\varphi_{Hi}(\mathbf{r}, t)$ ($i=x, y, z$)可能并不相等,为了以后书写方便,统一用 $\varphi_{Ei}(\mathbf{r})$ 表示初始相位,即选择 $\varphi_i(\mathbf{r}) = \varphi_{Ei}(\mathbf{r})$,而将 $\varphi_{Hi}(\mathbf{r})$ 写成 $\varphi_{Hi}(\mathbf{r}) = \varphi_i(\mathbf{r}) + \Delta_{EH}\varphi$ 。今后会看到这个 $\Delta_{EH}\varphi$ 当光波在线性介质中传播时保持不变。

将 $\cos[\omega t + \varphi(\mathbf{r})] = [e^{i[\omega t + \varphi(\mathbf{r})]} + e^{-i[\omega t + \varphi(\mathbf{r})]}] / 2$ 代入式(1.1.1-2)后,式中 e, e_* 表示这一项是求和号前面一项的共轭, $E_i(\mathbf{r}, t)$ 分为两支,一支是 $E_i(\mathbf{r}) e^{i[\omega t + \varphi_i(\mathbf{r})]} / 2$,另一支是 $E_i(\mathbf{r}) e^{-i[\omega t + \varphi_i(\mathbf{r})]} / 2$ 。在线性系统中,只需要研究其中的一支就可以了,于是取其中一支(负频率部分)为

$$E_i(\mathbf{r}) e^{-i[\omega t + \varphi_i(\mathbf{r})]} / 2 = E_i(\mathbf{r}) e^{-i\varphi_i(\mathbf{r})} e^{-i\omega t} / 2 \quad (1.1.1-3)$$

由于在线性系统中,频率是不变的,于是可以把它的频率项 $e^{-i\omega t}$ 不考虑,而且系数 $1/2$ 在线性系统中也不会变化,于是定义频域复振幅为

$$\dot{E}_i(\mathbf{r}, \omega) = E_i(\mathbf{r}) e^{-i\varphi_i(\mathbf{r})} \quad (1.1.1-4)$$

到了式(1.1.1-4),首先注意到式中不再出现时间变量 t 了,其次它也不再是一个实数了,它有大小和相位,所以称式(1.1.1-4)为光场的频域表达式的复振幅。复振幅这个术语有点别扭,因为我们知道描述正弦振动的三个要素(频率、振幅和相位)都是实数,这里出现一个“复振幅”就违背了原先的知识。请读者习惯这种打破过去已有概念的做法。其实,复振幅就是将振幅和相位组合在一起的一个物理量。这里还要特别说明,频域表达式是一定不含时间参数 t 的。

这时,将三个分量的频域表达式合写在一起,整体的公式变为

$$\dot{\mathbf{E}}(\mathbf{r}, \omega) = \sum_{i=1}^3 \dot{E}_i(\mathbf{r}, \omega) \hat{\mathbf{e}}_i = \sum_{i=1}^3 E_i(\mathbf{r}) e^{-i\varphi_i(\mathbf{r})} \hat{\mathbf{e}}_i \quad (1.1.1-5)$$

式中, $i=1, 2, 3$ 分别对应 $i=x, y, z$, $\hat{\mathbf{e}}_i$ 表示3个方向的单位矢量,并称 $\dot{\mathbf{E}}(\mathbf{r}, \omega)$ 为频域复矢量。这样,就得到了光场的频域表达式。

但是,在使用式(1.1.1-4)或者式(1.1.1-5)时,我们觉得很不方便,尤其是求和号是一个令人不舒服的符号,因为在做乘法时会出现很多项;而且,虽然这些公式已经完全描述了光场的全部特性,但是,概念还不够清楚。为此,将对它们做进一步的变形,从而引出整体的幅度、平均相位以及偏振态的概念。

经过严格的证明(附录 1-1),可以得出

$$\dot{\mathbf{E}}(\mathbf{r}, \omega) = E(\mathbf{r}, \omega) e^{i\varphi(\mathbf{r}, \omega)} \hat{\mathbf{e}}(\mathbf{r}, \omega) \quad (1.1.1-6)$$

式中: $E(\mathbf{r}, \omega) = \sqrt{\sum_{i=1}^3 |\dot{E}_i(\mathbf{r}, \omega)|^2}$, 称为频域复矢量的幅度, 请注意它是一个非负的实数; $\varphi(\mathbf{r}, \omega)$ 称为频域复矢量的相位, $\varphi = (\varphi_{E_x} + \varphi_{E_y} + \varphi_{E_z})/3$, 是三个分量相位的平均值, 也是一个实数, 可以为正或者为负; 最后一项 $\hat{\mathbf{e}}(\mathbf{r}, \omega)$ 是一个单位复矢量(模值为 1), 它与各个分量的关系为

$$\hat{\mathbf{e}}(\mathbf{r}, \omega) = e^{i\Delta\varphi_x(\mathbf{r}, \omega)} \cos\alpha \hat{\mathbf{x}} + e^{i\Delta\varphi_y(\mathbf{r}, \omega)} \cos\beta \hat{\mathbf{y}} + e^{i\Delta\varphi_z(\mathbf{r}, \omega)} \cos\gamma \hat{\mathbf{z}} \quad (1.1.1-7)$$

式中, $\cos\alpha, \cos\beta, \cos\gamma$ 为这个矢量的方向余弦, 分别为

$$\begin{cases} \cos\alpha = |\dot{E}_x(\mathbf{r}, \omega)| / E(\mathbf{r}, \omega) \\ \cos\beta = |\dot{E}_y(\mathbf{r}, \omega)| / E(\mathbf{r}, \omega) \\ \cos\gamma = |\dot{E}_z(\mathbf{r}, \omega)| / E(\mathbf{r}, \omega) \end{cases} \quad (1.1.1-8)$$

而 3 个相位分别为

$$\begin{cases} \Delta\varphi_x = (2\varphi_x - \varphi_y - \varphi_z)/3 \\ \Delta\varphi_y = (2\varphi_y - \varphi_x - \varphi_z)/3 \\ \Delta\varphi_z = (2\varphi_z - \varphi_x - \varphi_y)/3 \end{cases} \quad (1.1.1-9)$$

我们进一步将看到, 这个单位复矢量是光场对应的 $\hat{\mathbf{e}}(\mathbf{r}, \omega)$ 偏振态。

除电场强度可以用一个复矢量描述它的频域量, 磁场强度也可以表达为这种形式。此时, 为了和电场强度相区别, 我们对于相应的量加了一个下标 H , 而且对应公式的编号与电场的相同, 只是在末尾加了一个 H 。

对应于式(1.1.1-6), 对于磁场强度有

$$\dot{\mathbf{H}}(\mathbf{r}, \omega) = H(\mathbf{r}, \omega) e^{i\varphi_H(\mathbf{r}, \omega)} \hat{\mathbf{h}}(\mathbf{r}, \omega) \quad (1.1.1-6H)$$

式中, $H(\mathbf{r}, \omega) = \sqrt{\sum_{i=1}^3 |\dot{H}_i(\mathbf{r}, \omega)|^2}$, 称为频域复矢量的幅度, 请注意它是一个非负的实数; $\varphi_H(\mathbf{r}, \omega)$ 称为频域复矢量的相位, $\varphi_H = (\varphi_{H_x} + \varphi_{H_y} + \varphi_{H_z})/3$, 它是三个分量相位的平均值, 也是一个实数, 可以为正或者为负; 最后一项 $\hat{\mathbf{h}}(\mathbf{r}, \omega)$ 是一个单位复矢量(模值为 1), 它与各个分量的关系为

$$\hat{\mathbf{h}}(\mathbf{r}, t) = e^{i\Delta\varphi_{Hx}(\mathbf{r}, t)} \cos\alpha_H \hat{\mathbf{x}} + e^{i\Delta\varphi_{Hy}(\mathbf{r}, t)} \cos\beta_H \hat{\mathbf{y}} + e^{i\Delta\varphi_{Hz}(\mathbf{r}, t)} \cos\gamma_H \hat{\mathbf{z}} \quad (1.1.1-7H)$$

式中, $\cos\alpha_H$ 、 $\cos\beta_H$ 、 $\cos\gamma_H$ 为这个矢量的方向余弦, 分别为

$$\begin{cases} \cos\alpha_H = |\dot{H}_x(\mathbf{r}, \omega)| / H(\mathbf{r}, \omega) \\ \cos\beta_H = |\dot{H}_y(\mathbf{r}, \omega)| / H(\mathbf{r}, \omega) \\ \cos\gamma_H = |\dot{H}_z(\mathbf{r}, \omega)| / H(\mathbf{r}, \omega) \end{cases} \quad (1.1.1-8H)$$

而三个相位分别为

$$\begin{cases} \Delta\varphi_{Hx} = (2\varphi_{Hx} - \varphi_{Hy} - \varphi_{Hz})/3 \\ \Delta\varphi_{Hy} = (2\varphi_{Hy} - \varphi_{Hx} - \varphi_{Hz})/3 \\ \Delta\varphi_{Hz} = (2\varphi_{Hz} - \varphi_{Hx} - \varphi_{Hy})/3 \end{cases} \quad (1.1.1-9H)$$

同理, 这个单位复矢量是光场对应的 $\hat{\mathbf{h}}(\mathbf{r}, \omega)$ 偏振态。

对于光场的频域量, 可以使用光谱仪进行观察, 因为光功率正比于 $E^2(\mathbf{r}, \omega)$, 因此可以通过测量光功率的功率谱密度 $P(\omega)$ 获得大体上的结果, 它对应于 $|\dot{\mathbf{E}}(\omega)|^2$; 但是到目前为止, 仍然无法测量直接 $\dot{\mathbf{E}}(\omega)$ 的相位。后面还会讲到光场的时域描述, 即 $\dot{\mathbf{E}}(t)$ 。同样对于时域量, 也只能通过光示波器观察它的时域波形 $P(t)$, 相当于测定了 $|\dot{\mathbf{E}}(t)|^2$, 同样也测不出时域量 $\dot{\mathbf{E}}(t)$ 的相位。即使是利用频域量与时域量对应的傅里叶变换关系, 仍然解不出这两个相位。还有一点需要提醒读者, 光谱仪由于受到分辨率的限制, 不能测定很窄的光谱。目前的分辨率水平大约为皮米量级, 实际的光源已经达到千赫兹的水平, 需要用相干的方法才能测定。

有一种观点认为: 宽频光对应于窄脉冲, 这其实是不对的。尽管窄脉冲光的频带必然很宽, 但是反过来是不成立的。只有当宽频光的各个频率分量满足一定相位关系后, 才有可能在时域中表现为窄脉冲。不能笼统地说, 宽频光一定对应时域的窄脉冲。

图 1.1.1-1 是利用光谱仪记录下的某个光谱。当用示波器观察一个质量较好光源的输出波形时, 会看到一个几乎是直线的波形, 所以, 用光强随时间的变化来解释图 1.1.1-1 的光谱, 理论上是说不通的, 那么, 这个谱宽的成因是什么呢? 其实它是连续光的相位抖动引起的, 光源的线宽越窄, 说明其相位噪声越小。

最后, 应该看一下电场的频域复矢量与磁场的频域复矢量的关系, 这个关系是由麦克斯韦方程决定的, 将在 1.1.2 节讨论。

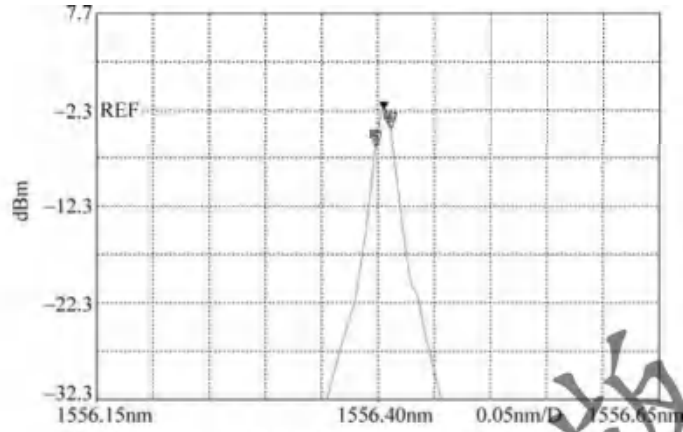


图 1.1.1-1 利用光谱仪记录的某个激光器的光谱

【附录】 1-1 复矢量的推导。

由式(1.1.1-5)可以得出

$$\begin{aligned}\dot{\mathbf{E}}(\mathbf{r}, \omega) &= E_x(\mathbf{r})e^{-i\varphi_x(\mathbf{r})}\hat{\mathbf{e}}_x + E_y(\mathbf{r})e^{-i\varphi_y(\mathbf{r})}\hat{\mathbf{e}}_y + E_z(\mathbf{r})e^{-i\varphi_z(\mathbf{r})}\hat{\mathbf{e}}_z \\ &= \sqrt{\sum_{i=1}^3 |\dot{E}_i(\mathbf{r}, \omega)|^2} \left\{ \frac{E_x(\mathbf{r})}{\sqrt{\sum_{i=1}^3 |\dot{E}_i(\mathbf{r}, \omega)|^2}} e^{-i\varphi_x(\mathbf{r})}\hat{\mathbf{e}}_x + \right. \\ &\quad \left. \frac{E_y(\mathbf{r})}{\sqrt{\sum_{i=1}^3 |\dot{E}_i(\mathbf{r}, \omega)|^2}} e^{-i\varphi_y(\mathbf{r})}\hat{\mathbf{e}}_y + \frac{E_z(\mathbf{r})}{\sqrt{\sum_{i=1}^3 |\dot{E}_i(\mathbf{r}, \omega)|^2}} e^{-i\varphi_z(\mathbf{r})}\hat{\mathbf{e}}_z \right\}\end{aligned}\quad (1.1.1-10)$$

定义 $E(\mathbf{r}, \omega) = \sqrt{\sum_{i=1}^3 |\dot{E}_i(\mathbf{r}, \omega)|^2}$, 且

$$\begin{cases} \cos\alpha = |\dot{E}_x(\mathbf{r}, \omega)| / E(\mathbf{r}, \omega) \\ \cos\beta = |\dot{E}_y(\mathbf{r}, \omega)| / E(\mathbf{r}, \omega) \\ \cos\gamma = |\dot{E}_z(\mathbf{r}, \omega)| / E(\mathbf{r}, \omega) \end{cases}\quad (1.1.1-11)$$

得到

$$\dot{\mathbf{E}}(\mathbf{r}, \omega) = E_x(\mathbf{r})e^{-i\varphi_x(\mathbf{r})}\hat{\mathbf{e}}_x + E_y(\mathbf{r})e^{-i\varphi_y(\mathbf{r})}\hat{\mathbf{e}}_y + E_z(\mathbf{r})e^{-i\varphi_z(\mathbf{r})}\hat{\mathbf{e}}_z$$

$$\begin{aligned}
&= E(\mathbf{r}, \omega) [\cos\alpha e^{-i\varphi_x(\mathbf{r})} \hat{\mathbf{e}}_x + \cos\beta e^{-i\varphi_y(\mathbf{r})} \hat{\mathbf{e}}_y + \cos\gamma e^{-i\varphi_z(\mathbf{r})} \hat{\mathbf{e}}_z] \\
&= E(\mathbf{r}, \omega) e^{-i\frac{\varphi_x + \varphi_y + \varphi_z}{3}} \left\{ \cos\alpha e^{-i\left[\varphi_x(\mathbf{r}) - \frac{\varphi_x + \varphi_y + \varphi_z}{3}\right]} \hat{\mathbf{e}}_x + \right. \\
&\quad \left. \cos\beta e^{-i\left[\varphi_y(\mathbf{r}) - \frac{\varphi_x + \varphi_y + \varphi_z}{3}\right]} \hat{\mathbf{e}}_y + \cos\gamma e^{-i\left[\varphi_z(\mathbf{r}) - \frac{\varphi_x + \varphi_y + \varphi_z}{3}\right]} \hat{\mathbf{e}}_z \right\} \quad (1.1.1-12)
\end{aligned}$$

定义 $\varphi = (\varphi_x + \varphi_y + \varphi_z)/3$ 和

$$\begin{cases} \Delta\varphi_x = (2\varphi_x - \varphi_y - \varphi_z)/3 \\ \Delta\varphi_y = (2\varphi_y - \varphi_x - \varphi_z)/3 \\ \Delta\varphi_z = (2\varphi_z - \varphi_x - \varphi_y)/3 \end{cases} \quad (1.1.1-9)$$

代入式(1.1.1-12), 得到

$$\hat{\mathbf{E}}(\mathbf{r}, \omega) = E(\mathbf{r}, \omega) e^{-i\varphi} [\cos\alpha e^{-i\Delta\varphi_x(\mathbf{r})} \hat{\mathbf{e}}_x + \cos\beta e^{-i\Delta\varphi_y(\mathbf{r})} \hat{\mathbf{e}}_y + \cos\gamma e^{-i\Delta\varphi_z(\mathbf{r})} \hat{\mathbf{e}}_z] \quad (1.1.1-13)$$

再定义单位复矢量

$$\hat{\mathbf{e}}(\mathbf{r}, \omega) = [\cos\alpha e^{-i\Delta\varphi_x(\mathbf{r})} \hat{\mathbf{e}}_x + \cos\beta e^{-i\Delta\varphi_y(\mathbf{r})} \hat{\mathbf{e}}_y + \cos\gamma e^{-i\Delta\varphi_z(\mathbf{r})} \hat{\mathbf{e}}_z] \quad (1.1.1-14)$$

最终得到式(1.1.1-6)。

1.1.2 频域麦克斯韦方程与亥姆霍兹方程

1.1.1 节已经得到了光场在频域中的表达式, 它可以归结为幅度、相位和偏振态三个量。本节将导出在这个前提下的麦克斯韦方程和亥姆霍兹方程。

如前所述, 麦克斯韦方程是描述一切电磁现象的基本方程, 然而, 在不同的场合, 其具体形式是不同的。因此, 我们首先导出光频下的麦克斯韦方程, 即频域麦克斯韦方程, 然后进一步导出频域亥姆霍兹方程。这两组方程的差别在于, 麦克斯韦方程是联系同一位置、某一时刻的电场强度与磁场强度之间关系的方程, 而亥姆霍兹方程是关于电场强度或者磁场强度单一的方程。亥姆霍兹方程虽然是由麦克斯韦方程导出的, 但二者并不完全等价。因为, 由亥姆霍兹方程导不出麦克斯韦方程。所以, 在某些情况下, 还是需要从麦克斯韦方程出发。

1. 频域麦克斯韦方程

在介质中基本的麦克斯韦方程在时域是瞬态的、在空间域是局部的, 即某一时刻、某一位置上电场强度 $\mathbf{E}$ 和磁场强度 $\mathbf{H}$ 所应满足的普适方程(无论是否是光频)为

$$\begin{cases} \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \\ \nabla \cdot \mathbf{D} = \rho \\ \nabla \cdot \mathbf{B} = 0 \end{cases} \quad (1.1.2-1)$$

第一个方程是说,在电磁场中,电场强度 $\mathbf{E}$ 的场是一个有旋场,产生这个旋度的源是同一点上磁感应强度 $\mathbf{B}$ 在这点随时间的变化率,通俗地说是“磁场的变化引起电场”。第二个方程是说,在电磁场中,磁场强度 $\mathbf{H}$ 的场是一个有旋场,产生这个旋度的源是同一点上电位移矢量 $\mathbf{D}$ 在这点随时间的变化率和电荷的运动(电流密度矢量 $\mathbf{J}$),通俗地说是“电场的变化和电荷的运动引起磁场”。第三个方程是说,电位移矢量 $\mathbf{D}$ 的场是一个有源场,产生这个散度的源是电荷密度 ρ 。第四个方程是说,磁感应强度 $\mathbf{B}$ 的场是一个无源场,自然界中没有“磁荷”存在。在这四个方程中,前两个是基本的,后两个可利用旋度场的散度恒为零以及电荷不灭定律

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} \quad (1.1.2-2)$$

导出。

在大多数实际光场中,既不存在传导电流,也不存在空间电荷,因此式(1.1.2-1)可改写为

$$\begin{cases} \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \\ \nabla \cdot \mathbf{D} = 0 \\ \nabla \cdot \mathbf{B} = 0 \end{cases} \quad (1.1.2-3)$$

关于麦克斯韦方程中 $\mathbf{D}$ 与 $\mathbf{E}$ 、 $\mathbf{B}$ 与 $\mathbf{H}$ 的关系(又称物性方程),是由波导材料的性质决定的,为

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (1.1.2-4a)$$

$$\mathbf{B} = \mu_0 \mathbf{H} + \mathbf{M} \quad (1.1.2-4b)$$

式中, ϵ_0 为真空中的介电常数, μ_0 为真空中的导磁率, $\mathbf{P}$ 和 $\mathbf{M}$ 分别为电极化强度和磁化强度。

在光频下,介质都是无磁性介质,即 $\mathbf{M} = \mathbf{0}$,于是

$$\mathbf{B} = \mu_0 \mathbf{H} \quad (1.1.2-5)$$

但 $\mathbf{P}$ 与 $\mathbf{E}$ 之间的关系却可能是很复杂的, $\mathbf{P}$ 与 $\mathbf{E}$ 关系的每一个微小变化,都将导致光波导新的物理现象。这里,我们首先把注意力集中于线性、各向同性的时不变

光波导,这时

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} = \epsilon_0 \left\{ \mathbf{E} + \int_{-\infty}^{+\infty} x^{(1)}(t-t_1) \cdot \mathbf{E}(\mathbf{r}, t_1) dt_1 \right\} \quad (1.1.2-6)$$

这个公式的物理解释为,电位移矢量 $\mathbf{D}$ 是由真空中的电场强度 $\epsilon_0 \mathbf{E}$ 和介质在电场作用下产生的电极化矢量 $\mathbf{P}$ 两部分组成。也就是说,介质在电场的作用下被极化,这个极化场叠加在原有的场上,二者共同构成了新的场。当电场强度 $\mathbf{E}$ 不随时间变化时, $\mathbf{P}$ 与 $\mathbf{E}$ 呈矢量比例关系(也就是大小呈比例关系,方向呈固定的旋转关系);但当电场强度 $\mathbf{E}$ 随时间变化时, $\mathbf{P}$ 与 $\mathbf{E}$ 呈卷积关系。这个关系可以用简单的因果律得到。因果律可以大致地概括为,一件事物的果应该是这个事物所有事前的因共同作用的结果。所以

$$\mathbf{P}(t) dt = \epsilon_0 f[\mathbf{E}(\mathbf{r}, t-t_1) dt_1] \quad (1.1.2-7)$$

式中, $f[\cdot]$ 是一个因果函数, t_1 是事前的某个时间,它的作用效果一直延续到 $t-t_1$ 时间。对于线性的因果关系, $f[\cdot]$ 表现为一个线性变换,对于矢量而言,它是一个 3×3 的矩阵,于是

$$\mathbf{P}(t) dt = \epsilon_0 x^{(1)}(t-t_1) \cdot \mathbf{E}(\mathbf{r}, t_1) dt_1 \quad (1.1.2-8)$$

总的累计效果为

$$\mathbf{P}(t) = \epsilon_0 \int_{-\infty}^{+\infty} x^{(1)}(t-t_1) \cdot \mathbf{E}(\mathbf{r}, t_1) dt_1 \quad (1.1.2-9)$$

考虑到事后的时间 $(t, +\infty)$ 区间内,作用的效果为 0,于是得到

$$\mathbf{P} = \epsilon_0 \int_{-\infty}^{+\infty} x^{(1)}(t-t_1) \cdot \mathbf{E}(\mathbf{r}, t_1) dt_1 \quad (1.1.2-10)$$

将它代入式(1.1.2-4a),再作一个变量替换,就得到式(1.1.2-6)。

式(1.1.2-6)表明, $\mathbf{P}$ 的变化要比 $\mathbf{E}$ 的变化滞后一定时间。只有当 $\mathbf{E}$ 变化很慢时,这种滞后效应才不显现出来,但光是频率极高的电磁波,这种滞后关系是很明显的。对式(1.1.2-6)进行傅里叶变换,可得

$$\mathbf{D}(\omega) = \hat{\epsilon}(\omega) \mathbf{E}(\omega) \quad (1.1.2-11)$$

因此,我们应注意,以往常用的表达式 $\mathbf{D} = \epsilon \mathbf{E}$,实际上是一个频域表达式,而不是时域表达式。一般来说,公式 $\mathbf{D}(t) = \epsilon \mathbf{E}(t)$ 是不正确的,只有在它们随时间的变化很慢时才成立。在式(1.1.2-11)中, $\mathbf{D}$ 、 $\mathbf{E}$ 、 ϵ 三个量均是频率(光频)的函数,都有大小、相位等。在实际应用中,常把 $\mathbf{D}(\omega)$ 、 $\mathbf{E}(\omega)$ 和 $\epsilon(\omega)$ 中的频率 ω 隐去,写成 $\mathbf{D}$ 、 $\mathbf{E}$ 、 ϵ ,我们只要记住它们实际上都是频域量就可以了。如同 1.1.1 节中对于电场强度和磁场强度分解为频域复矢量一样,对于电位移矢量,同样可以求得它的频域表达式为

$$\hat{\mathbf{D}} = D(\mathbf{r}, \omega) e^{-i\varphi_D(\mathbf{r}, \omega)} \hat{\mathbf{d}}(\mathbf{r}, \omega) \quad (1.1.2-12)$$

于是,式(1.1.2-11)可改写为

$$\hat{\mathbf{D}}(\omega) = \hat{\epsilon}(\omega) \hat{\mathbf{E}}(\omega) \quad (1.1.2-13)$$

式中, $\hat{\epsilon}(\omega)$ 是式(1.1.2-6)在卷积所对应的频域量, 通常是一个复数, 实部表示相移部分, 虚部表示损耗, 参见 11.1.1 节。

现在, 将 1.1.1 节中光场的频域表达式(1.1.1-6)再乘以 $e^{-i\omega t}$ 代入麦克斯韦方程(1.1.2-1), 进行微分运算后, 然后再消去 $e^{-i\omega t}$, 得到

$$\nabla \times \dot{\mathbf{E}} = i\omega\mu_0 \dot{\mathbf{H}} \quad (1.1.2-14a)$$

$$\nabla \times \dot{\mathbf{H}} = -i\omega\dot{\mathbf{D}} \quad (1.1.2-14b)$$

$$\nabla \cdot \dot{\mathbf{D}} = 0 \quad (1.1.2-14c)$$

$$\nabla \cdot \dot{\mathbf{H}} = 0 \quad (1.1.2-14d)$$

式(1.1.2-14)就是用频域量表示的麦克斯韦方程。

2. 介质的性质

在式(1.1.2-14)中, 除了电场强度和磁场强度两个频域量外, 还有一个变量——电位移矢量。因此, 使用起来不方便。本小节的工作, 是试图把电位移矢量从方程组中消去, 这就涉及介质的性质。

最简单的介质是线性、各向同性且时不变的介质, 也就是用式(1.1.2-13)所描述的那种介质。将式(1.1.2-13)代入式(1.1.2-14c), 利用 $\nabla \cdot (\epsilon \mathbf{E}) = \nabla \epsilon \cdot \mathbf{E} + \epsilon \nabla \cdot \mathbf{E} = 0$, 得到

$$\nabla \cdot \mathbf{E} = \frac{-\nabla \epsilon}{\epsilon} \cdot \mathbf{E} \quad (1.1.2-15)$$

此式有明确的物理意义, 即波导中介质分布的任何不均匀性, 在电场 $\mathbf{E}$ 的作用下, 将使 $\mathbf{E}$ 成为有源场, 尽管此处并无空间电荷 ρ 。它的物理解释为: 介质分布的不均匀性, 导致极化电荷分布得不均匀, 出现微观剩余电荷, 表现为有源场。将它代入式(1.1.2-1), 就得到了线性时不变介质的频域麦克斯韦方程。

$$\nabla \times \dot{\mathbf{E}} = i\omega\mu_0 \dot{\mathbf{H}} \quad (1.1.2-16a)$$

$$\nabla \times \dot{\mathbf{H}} = -i\omega\hat{\epsilon}\dot{\mathbf{E}} \quad (1.1.2-16b)$$

$$\nabla \cdot \dot{\mathbf{E}} = -\frac{\nabla \hat{\epsilon}}{\hat{\epsilon}} \cdot \dot{\mathbf{E}} \quad (1.1.2-16c)$$

$$\nabla \cdot \dot{\mathbf{H}} = 0 \quad (1.1.2-16d)$$

简单介绍一下各向异性介质的问题。描述介质性质的式(1.1.2-13)仅适用于各向同性的线性介质, 也就是说, $\hat{\epsilon}(\omega)$ 的值与注入电场的方向无关。但是很多材料的折射率(或者介电常数)与注入电场的方向有关, 不同方向的电场分量有不同的介电常数, 这种介电常数与电场强度的方向有关的现象, 称为各向异性。这时就一般情况而言, 介电常数是一个 3×3 的矩阵, 这个矩阵可以称为介电常数矩阵, 或

者介电常数张量,简称为介电张量。注意这个矩阵的每个分量均是频率(光频)的函数,都有大小、相位等,即

$$\hat{\epsilon}(\omega) = \begin{bmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz} \\ \epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz} \end{bmatrix} (\omega) \quad (1.1.2-17)$$

于是

$$D(\omega) = \begin{bmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz} \\ \epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz} \end{bmatrix} E(\omega) \quad (1.1.2-18)$$

各向异性通常在晶体内发生,而晶体的内部结构有一定的对称性,这种对称性导致矩阵 $\hat{\epsilon}(\omega)$ 也具有一定的对称性。根据矩阵理论,具有一定对称性的矩阵都可以对角化。数学上的对角化对应物理上的坐标变换。也就是说,式(1.1.2-17)和式(1.1.2-18)经过一定的坐标变换,可以得到

$$\hat{\epsilon} = \begin{bmatrix} \epsilon_x & & \\ & \epsilon_y & \\ & & \epsilon_z \end{bmatrix} = \epsilon_0 \hat{\epsilon}_r \quad (1.1.2-19)$$

$$D(\omega) = \begin{bmatrix} \epsilon_x & & \\ & \epsilon_y & \\ & & \epsilon_z \end{bmatrix} E(\omega) = \epsilon_0 \hat{\epsilon}_r E(\omega) \quad (1.1.2-20)$$

满足矩阵对角化条件的坐标系,称为主轴坐标系。

光在各向异性材料中的传播,表现出奇特的光学现象,其中一种重要的光学现象就是所谓双折射现象。当我们透过双折射材料(如方解石)看一张纸上的一个字时,将看到两个字的像,好像有两条折射线一般,这就是双折射现象,如图1.1.2-1所示。关于各向异性材料中的光学现象,将在第9章讨论。



图 1.1.2-1 透过方解石的双折射现象

(请扫 II 页二维码看彩图)

关于光频下各向异性材料的麦克斯韦方程,不能简单地将 $\dot{\mathbf{D}} = \hat{\epsilon} \dot{\mathbf{E}}$ 代入式(1.1.2-14c)中,这是因为数学上不能得到 $\nabla \cdot (\hat{\epsilon} \dot{\mathbf{E}}) = \nabla \hat{\epsilon} \cdot \dot{\mathbf{E}} + \hat{\epsilon} \nabla \cdot \dot{\mathbf{E}}$ 。很明显,这个式子的左边是一个数量,右边有如下困难:一个矩阵 $\hat{\epsilon}$ 的梯度没有定义;而 $\hat{\epsilon} \nabla \cdot \mathbf{E}$ 是一个矩阵,因此,左右两边是不可能相等的。关于光频下各向异性材料的麦克斯韦方程的问题,我们将留在第9章解决。

3. 频域亥姆霍兹方程

麦克斯韦方程组中的每一个方程,都包含电场与磁场两个量,它们互相关联,使用起来很不方便。为了获得电场与磁场各自单独满足的方程,对方程组(1.1.2-16)进行简单的数学演算,即可将 $\mathbf{E}$ 和 $\mathbf{H}$ 互相关联的方程转化为各自独立的方程,例如对式(1.1.2-16a),利用

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} \quad (1.1.2-21)$$

立即可得亥姆霍兹方程

$$\nabla^2 \dot{\mathbf{E}} + k_0^2 n^2 \dot{\mathbf{E}} + \nabla \left(\dot{\mathbf{E}} \cdot \frac{\nabla \epsilon}{\epsilon} \right) = 0 \quad (1.1.2-22a)$$

$$\nabla^2 \dot{\mathbf{H}} + k_0^2 n^2 \dot{\mathbf{H}} + \frac{\nabla \epsilon}{\epsilon} \times (\nabla \times \dot{\mathbf{H}}) = 0 \quad (1.1.2-22b)$$

式中, $k_0 = 2\pi/\lambda$ 为真空中的波数, λ 为波长; $n^2 = \epsilon/\epsilon_0$ 。顺便说明,今后介电常数 $\hat{\epsilon}$ 的虚部常常被忽略掉, $\hat{\epsilon} \approx \epsilon$, 才有上述结果。绝大多数文献并不明确说式(1.1.2-22)是一个频域表达式,而笼统地写成

$$\begin{cases} \nabla^2 \mathbf{E} + k_0^2 n^2 \mathbf{E} + \nabla \left(\mathbf{E} \cdot \frac{\nabla \epsilon}{\epsilon} \right) = 0 \\ \nabla^2 \mathbf{H} + k_0^2 n^2 \mathbf{H} + \frac{\nabla \epsilon}{\epsilon} \times (\nabla \times \mathbf{H}) = 0 \end{cases} \quad (1.1.2-23)$$

这其实是不准确的。如果不强调一下是频域表达式,以为对于时域的 $\mathbf{E}(t)$ 也可以使用,那就可能引起错误。

由亥姆霍兹方程(1.1.2-22)可以看出,方程的左边包括齐次部分 $(\nabla^2 + k_0^2 n^2) \cdot \begin{pmatrix} \dot{\mathbf{E}} \\ \dot{\mathbf{H}} \end{pmatrix}$ 和非齐次部分。而 $\nabla \epsilon$ 是否为零是该方程是否为齐次的关键。如果在所考虑的那一部分光波导中,介质为均匀分布, $\nabla \epsilon = 0$, 或近似均匀分布, $\nabla \epsilon / \epsilon \rightarrow 0$, 那么该方程就转化为齐次方程

$$(\nabla^2 + k_0^2 n^2) \begin{pmatrix} \dot{\mathbf{E}} \\ \dot{\mathbf{H}} \end{pmatrix} = 0 \quad (1.1.2-24)$$

齐次方程的特点是可以使用分离变量法,把各分量区分开,而非齐次方程则不然。所以,我们依据折射率分布的均匀性对光波导进行分类,实质上是看它的光场满足